ABSTRACT

Disclosed are a computer-readable code, system and method for comparing a target concept, invention, or event with each of a plurality of texts. Each of a plurality of non-generic words and optionally, words groups

5   characterizing the target concept, invention, or event, is selected as a vector term if the term has an above-threshold selectivity value in at least one library of texts in a field, where the selectivity value of a term is a measure of the field-specificity of that term. There is then determined, for each of the plurality of texts, a match score related to the number of vector terms present in or derived from that text that

10   match those in the target concept, invention, or event. Texts having the highest match scores are selected.